

The inexact projected gradient method for quasiconvex vector functions

J. Y. Bello Cruz ^{*} G. C. Bento[†] G. Bouza Allende[‡] R. F. B. Costa[‡]

December 6, 2012

Abstract

In this work, we consider the inexact projected gradient-like method for solving smooth constrained vector optimization problems. Basically, assuming that the objective function of the problem is K -quasiconvex, we prove global convergence of any sequence produced by the method to a critical Pareto point.

Keywords: Gradient-like method · vector optimization · K -quasiconvexity.

Mathematical Subject Classification (2010): 90C26, 90C29, 90C31.

1 Introduction

In many applications, such as engineers, statistics and design problems, several objective functions have to be minimized simultaneously; see [8, 12, 13, 16, 22, 31, 33]. Sometimes the preferences are not described by the Pareto cone, and even in this case the minimizer of the objective functions may be different. So, the concept of optimality has to be replaced by the concept of efficiency.

A popular strategy for solving vector optimization problems is the scalarization approach. The most widely used scalarization technique is the weighting method; see [20, 28]. Basically, one minimizes a linear positive combination of the objectives, where the vector of “weights” is not known a priori. This procedure may lead to unbounded numerical problems, which, therefore, may lack minimizers; see [20, 28, 29]. Another disadvantage of this approach is that the choice of the parameters is not known in advance, leaving the modeler and the decision-maker with the burden of choosing them.

The class of quasiconvex vectorial functions has many applications in the real life problems. For instance, usually utility functions in economy are quasiconcave functions; see [23, 24]. For quasiconvex vector optimization problem the weighting method has another weakness: the objective function of the scalar problem may not be quasiconvex. So, other approaches must be investigated. Recently, the gradient method for multiobjective optimization problems was proposed in Fliege and Svaiter [15]. Since then, it has been considered in more general settings, for instance, for vector optimization problems; see Graña-Drummond and Svaiter [21], and for

^{*}IME, Universidade Federal de Goiás, Goiânia, GO 74001-970, BR. Tel.: +55 62 3521 1418; fax: +55 62 3521 1180. yunier@impa.br. **Corresponding author.**

[†]IME, Universidade Federal de Goiás, Goiânia, BR.

[‡]Departamento de Matemática Aplicada, Facultad de Matemática y Computación, Universidad de La Habana, Habana, Cuba.

constrained vector optimization; see [18, 19]. As far as we know, the first results of full convergence of the exact gradient method for quasiconvex multicriteria optimization appeared in Bento et al. [3] and Bello et al. [1]. The convergence of the inexact variant as well as its links with a psychological model of self regulation has been studied in Bento et al. [4]. Although other classical methods for solving scalar optimization problems have been extended for the setting of vector case; see [5, 9, 11, 14, 15, 19], in the present paper we study the inexact projected gradient method, proposed by Fukuda and Graña-Drummond [17]. We prove that the sequence generated by it converges to a stationary point of the vector optimization problem in the quasiconvex case.

This article is organized as follows. In Section 2, we outline some basic definitions, assumptions and preliminary materials. Then, in Section 3, we present the inexact projected gradient method for vectorial optimization. Finally, Section 4 contains the convergence analysis of the method.

2 Basic definitions and preliminary material

In this section, we present the vector optimization problem as well as some definitions, notations and basic properties, which are used throughout of this paper. For more details; see, for instance, [2, 17, 21].

Let $K \subset \mathbb{R}^m$ be a nonempty closed, convex and pointed cone (i.e., $K \cap -K = \{0\}$). The partial order “ \preceq_K ” induced by K in \mathbb{R}^m is defined as follows: $u \preceq_K v$ if, and only if, $v - u \in K$. Assuming that K has nonempty interior, i.e. $\text{int}(K) \neq \emptyset$, we say that $u \prec_K v$ if, and only if, $v - u \in \text{int}(K)$.

Important properties only hold if the partial order is directed, i.e. for all $y_1, y_2 \in \mathbb{R}^m$, there exists $z \in \mathbb{R}^m$ such that $y_1 \preceq_K z$ and $y_2 \preceq_K z$. Next proposition shows that in our case this condition holds.

Proposition 1. *If $\text{int}(K) \neq \emptyset$, then $(\mathbb{R}^m, \preceq_K)$ is directed.*

Proof. See [2, Remark 2.2]. □

Given $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ and $C \subseteq \mathbb{R}^n$ a nonempty set, we consider the problem of finding a weakly efficient point of F in C , i.e., a point $x^* \in C$ such that there exists no other $x \in C$ with $F(x) \prec_K F(x^*)$. We denote this constrained problem as

$$\begin{aligned} & \min_K F(x) \\ & \text{s.t. } x \in C. \end{aligned} \tag{1}$$

From now on $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is continuously differentiable and $C \subseteq \mathbb{R}^n$ is a nonempty, closed and convex set. In this particular framework, the necessary condition proven in Chen and Jahn [10] leads to the following: if $x^* \in C$ is weakly efficient point, then x^* is a stationary point, i.e.,

$$-\text{int}(K) \cap J_F(x^*)(C - x^*) = \emptyset, \tag{2}$$

where $J_F(x)$ is the Jacobian matrix of F at x and $C - x^* = \{y - x^* : y \in C\}$. In fact, x^* is a stationary point if, and only if, for all $v \in C - x^*$, we have $J_F(x^*)v \not\prec_K 0$. As reported in Chen and Jahn [10], in general, condition (2) is not sufficient for a point to be a weakly efficient solution of (1).

Now we present the concept of descent direction.

Definition 1. *We say that $v \in \mathbb{R}^n$ is a descent direction at $x \in C$, if $v \in C - x$ and $J_F(x)v \prec_K 0$.*

For finding a decreasing direction we use the positive polar cone of K , i.e, the set

$$K^* := \{y \in \mathbb{R}^m; \langle y, x \rangle \geq 0 \text{ for all } x \in K\}.$$

Due to the fact that $K \subset \mathbb{R}^m$ is a convex and closed cone, K^* is also convex and closed. Since we only need to find v such that $J_F(\bar{x})v \prec_K 0$, the analysis can be reduced to a compact set of normalized generators of K^* . Recall that G is a set of normalized generators of K^* if $G \subset \{y \in K^*; \|y\| = 1\}$ is compact and K^* is the cone generated by its convex hull. Such a set always exists; one can take for example $G = \{y \in K^*; \|y\| = 1\}$, but in general it is possible to consider much smaller sets; see [26, 27, 29]. In the multiobjective case, $K = \mathbb{R}_+^m$, G can be taken as the canonical basis of \mathbb{R}^m because $(\mathbb{R}_+^m)^* = \mathbb{R}_+^m$. A good choice of G can be obtained in terms of extreme directions. Let us present its definition.

Definition 2. Let $K \subset \mathbb{R}^m$ be a convex and closed cone. We say that $d \in \text{extd}(K)$, i.e. d is a extreme direction of K , if and only if $d \in K \setminus \{0\}$ and for all $d_1, d_2 \in K$ such that $d = d_1 + d_2$ we have $d_1, d_2 \in \mathbb{R}_+ d$.

The following result holds.

Proposition 2. If $\text{int}(K) \neq \emptyset$, then K^* is the conic hull of closed convex hull of $\text{extd}(K^*)$, i.e.,

$$K^* = \text{co}(\text{conv}(\text{extd}(K^*))).$$

Proof. See [2, Remark 3.2]. □

For an equivalent description of the relation $y \in K$, we consider the following auxiliary function. Let $\varphi : \mathbb{R}^m \rightarrow \mathbb{R}$ be the function defined by

$$\varphi(y) := \max \langle y, \omega \rangle \quad \text{s.t. } \omega \in G.$$

In terms of φ we have the following characterization of $-K$ and $-\text{int}(K)$; see Graña-Drummond and Svaiter [21],

$$-K = \{y \in \mathbb{R}^m : \varphi(y) \leq 0\} \quad \text{and} \quad -\text{int}(K) = \{y \in \mathbb{R}^m : \varphi(y) < 0\}.$$

Next we present some basic properties of φ .

Proposition 3.

- (i) The function φ is positively homogeneous of degree 1.
- (ii) $\varphi(y + z) \leq \varphi(y) + \varphi(z)$ and $\varphi(y) - \varphi(z) \leq \varphi(y - z)$ for all $y, z \in \mathbb{R}^m$.
- (iii) Given $y, z \in \mathbb{R}^m$, if $y \prec_K z$ ($y \preceq_K z$), then $\varphi(y) < \varphi(z)$ ($\varphi(y) \leq \varphi(z)$, respectively).
- (iv) The function $\varphi : \mathbb{R}^m \rightarrow \mathbb{R}$ is Lipschitz continuous with Lipschitz constant $L = 1$.

Proof. See [19, Proposition 2]. □

For a fixed $x \in C$, consider the following constrained optimization problem:

$$\begin{aligned} & \min h_x(v) \\ & \text{s.t. } v \in C - x \end{aligned} \tag{3}$$

where $h_x(v) := \widehat{\beta} \varphi(J_F(x)v) + \frac{\|v\|^2}{2}$, with $\widehat{\beta} > 0$.

We resume some property associated to Problem (3) in the following lemma.

Lemma 1. *The following statements hold:*

- (i) *The constrained optimization problem given in (3) has only one solution. Moreover, the vector v is the solution of (3) if and only if there exists $\omega \in \text{conv}(G)$ such that*

$$v = P_{C-x}(-\hat{\beta}J_F(x)^T\omega).$$

- (ii) *If x is a stationary point of F and v solves Problem (3), then $v = 0$ and 0 is the optimal value of $h_x(v)$.*

- (iii) *If $x \in \mathbb{R}^n$ is a non-stationary point of F and v is the solution of (3), then $v \neq 0$ and*

$$h_x(v) < 0.$$

In particular, v is a descent direction for F at x .

Proof. See [17, Proposition 4.1] and [19, Proposition 3]. □

In view of the previous lemma and (3), we define the following functions.

Definition 3. *The projected gradient direction function of F is defined as $v : C \rightarrow \mathbb{R}^n$, where $v(x)$ is the unique solution of Problem (3).*

The optimal value function associated to (3) is $\theta : C \rightarrow \mathbb{R}$, where $\theta(x) := h_x(v(x))$.

Remark 1. *This definition was proposed in Graña-Drummond and Iusem [19]. Note that, from Lemma 1(i), it follows that the projected gradient direction function of F is well defined. It is interesting to observe that we are in fact extending the real-valued projected gradient direction. Indeed, for $m = 1$, since $K = \mathbb{R}_+$ and $G = \{1\}$, we retrieve the following identities:*

$$v(x) = P_{C-x}(-\hat{\beta}\nabla F(x)) = P_C(x - \hat{\beta}\nabla F(x)) - x.$$

Lemma 2. *The projected gradient direction function of F , $v(x)$ and the value function $\theta(x)$ are continuous.*

Proof. For the proof of the continuity of $v(x)$; see [18, Proposition 3.4]. The second part is a direct consequence of this fact. □

Now we consider the inexact case. Let us present the concept of approximate directions.

Definition 4. *Let $x \in C$ and $\sigma \in [0, 1)$. A vector $v \in C - x$ is a σ -approximate projected gradient direction at x if*

$$h_x(v) \leq (1 - \sigma)\theta(x).$$

Remark 2.

- (i) *Note that the exact direction $v = v(x)$ is always a σ -approximate at x for any $\sigma \in [0, 1)$. Moreover, $v(x)$ is the unique 0-approximate direction at x .*
- (ii) *Given a nonstationary point $x \in C$ and $\sigma \in [0, 1)$, a σ -approximate direction v is always a descent direction. Indeed, since $\sigma \in [0, 1)$, this follows from the definitions of the functions $\theta(\cdot)$ and $h_x(\cdot)$, combined with the Proposition 1(iii) and Definition 4.*

A particular class of σ -approximate directions for F at x is given by the directions $v \in \mathbb{R}^n$ which are *scalarization compatible* (or simply *s-compatible*), i.e., those $v \in \mathbb{R}^n$ such that there exists $\omega \in \text{conv}(G)$ with

$$v = P_{C-x}(-\widehat{\beta} J_F(x)^T \omega).$$

Note that ω determines a scalar function $g(x) := \langle \omega, F(x) \rangle$ whose projected gradient direction associated to the scalar-valued problem

$$\begin{aligned} & \min g(x) \\ & \text{s.t. } x \in C, \end{aligned}$$

coincides with v . This justifies the name previously attributed to the direction v ; see Fukuda and Graña-Drummond [17] for a discussion. From Lemma 1(i), it follows that the exact search direction $v(x)$ is *s-compatible*. In Fukuda and Graña-Drummond [17] the authors also presented a sufficient condition for *s-compatible* direction to be a σ -approximation direction for F at x . We end this section with some results on quasiconvexity.

Definition 5. *The vector function $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is said to be K -quasiconvex if for all $y \in \mathbb{R}^m$ the level set $L_F(y) = \{x \in \mathbb{R}^n : F(x) \preceq_K y\}$ is convex.*

In the scalar case we remind that a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is quasiconvex iff $f(\alpha x + (1 - \alpha)y) \leq \max\{f(x), f(y)\}$ for every $\alpha \in [0, 1]$ and $x, y \in \mathbb{R}^n$. Moreover under differentiability, quasiconvexity is equivalent to say that for each $x, y \in \mathbb{R}^n$, the inequality $f(x) \leq f(y)$ implies that $\langle \nabla f(y), x - y \rangle \leq 0$; see [30, Theorem 9.1.4].

The Definition 5 is milder than K -convexity. Recall that F is a K -convex function if the set $\{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m : F(x) \preceq_K y\}$ is convex.

Now we can give the following characterization of K -quasiconvexity in terms of scalar quasiconvexity:

Theorem 1. *Assume that $(\mathbb{R}^m, \preceq_K)$ is partially ordered. Then the following assertions are equivalent:*

(i) F is K -quasiconvex;

(ii) $\langle d, F(\cdot) \rangle : \mathbb{R}^n \rightarrow \mathbb{R}$ is quasiconvex for every extreme direction $d \in K^*$.

Proof. Since $\text{int}(K) \neq \emptyset$, it follows from Proposition 2 that K^* is the conic hull of the closed convex hull of $\text{extd}(K^*)$. Again, as $\text{int}(K) \neq \emptyset$, by Proposition 1, we have that $(\mathbb{R}^m, \preceq_K)$ is directed. Combining these two facts, the desired result, follows from [2, Theorem 3.1]. \square

3 Inexact projected gradient algorithm

This part is devoted to present the method and some basic properties of it. Let us first consider the following constants: $\widehat{\beta} > 0$, $\delta \in (0, 1)$, $\tau > 1$ and $\sigma \in [0, 1)$. The inexact projected gradient method is defined as follows.

Initialization: Take $x^0 \in C$.

Iterative step: Given x^k , compute a σ -approximate direction v^k at x^k .

If $h_{x^k}(v^k) = 0$, then stop. Otherwise compute

$$j(k) = \min \{j \in \mathbb{Z}_+ : F(x^k + \tau^{-j} v^k) \preceq_K F(x^k) + \delta \tau^{-j} J_F(x^k) v^k\}. \quad (4)$$

Set $t_k = \tau^{-j(k)}$ and $x^{k+1} = x^k + t_k v^k$.

The previous algorithm was proposed by Fukuda-Graña-Drummond in [17]. In particular, when $m = 1$ and $\sigma = 0$, it becomes the classical exact projected gradient method. In the inexact unconstrained case, we retrieve the steepest descent method introduced by Graña-Drummond and Svaiter in [21]. Note that this approach is a natural extension of the method proposed in Graña-Drummond and Iusem [19]. We observe that the stopping criterion $h_{x^k}(v^k) = 0$, implies the stationarity of x^k . Indeed, if $h_{x^k}(v^k) = 0$, then, by the definition of σ -approximation, $\theta(x^k) \geq 0$, but as $\theta(x^k) \leq 0$, so $\theta(x^k) = 0$, concluding that x^k is a stationary point. On the other hand, if x^k is stationary, then $\theta(x^k) = 0$, and therefore $h_{x^k}(v^k) = 0$.

As already explained in Remark 2(ii), if x^k is not a stationary point, then v^k is a descent direction, i.e., $J_F(x^k)v^k \prec_K 0$. Next proposition shows that the Armijo rule given in (4) is well defined.

Proposition 4. *Let $\delta \in (0, 1)$, $x \in C$ and let v be a descent direction. Then, there exists $\bar{\gamma} > 0$ fulfilling that $F(x + \gamma v) \prec_K F(x) + \delta \gamma J_F(x)v$ for all $\gamma \in (0, \bar{\gamma}]$.*

Proof. See [19, Proposition 1]. □

4 Convergence analysis

In this section, we show that every sequence produced by the inexact projected gradient method converges globally to a stationary point, under reasonable hypotheses. The novelty presented in this paper is that the convergence result only needs that the objective function is K -quasiconvex. As already remarked, if the method stops after a finite number of iterations, then it stops at a stationary point. From now on, we will assume that $\{x^k\}$, $\{v^k\}$, $\{t_k\}$ are the infinite sequences generated by the inexact projected gradient method. We begin with some known previous results.

Definition 6. *A sequence $\{y^k\} \subset \mathbb{R}^m$ is said K -bounded if there exists $z \in \mathbb{R}^m$ such that $z \preceq_K y^k$, for all k .*

Lemma 3. *If the sequence $\{F(x^k)\}$ is K -bounded, then we have*

$$\sum_{k=0}^{\infty} t_k |\langle \omega, J_F(x^k)v^k \rangle| < +\infty, \quad \omega \in \text{conv}(G).$$

Proof. See [17, Lemma 3.6]. □

Proposition 5. *The sequence generated by the inexact projected gradient method is feasible and $\{F(x^k)\}$ is a K -decreasing sequence.*

Proof. The feasibility of the sequence $\{x^k\}$ is a consequence of the definition of the method and the K -decreasing property follows from (4). □

Next, we deal with so called quasi-Fejér convergence and its proprieties, because the convergence analysis of the proposed method is based on this theory.

Definition 7. *A sequence $\{x^k\} \subset \mathbb{R}^n$ is said to be quasi-Fejér convergent to a set V , $V \neq \emptyset$, if and only if for each $x \in V$ there exists a summable sequence $\{\epsilon_k\} \subset \mathbb{R}_+$, such that*

$$\|x^{k+1} - x\|^2 \leq \|x^k - x\|^2 + \epsilon_k, \quad k = 0, 1, \dots$$

This definition was originated in Browder [6] and has been further elaborated in Iusem et al. [25]. A useful result on quasi-Fejér sequences is the following.

Lemma 4. *If $\{x^k\} \subset \mathbb{R}^n$ is quasi-Fejér convergent to some set $V \neq \emptyset$, then:*

- i) *The sequence $\{x^k\}$ is bounded;*
- ii) *If, furthermore, an accumulation point x of the sequence $\{x^k\}$ belongs to V , then $\{x^k\}$ converges to x .*

Proof. See [7, Theorem 1]. □

Under differentiability, the main result on the convergence of the inexact projected gradient method is the following.

Proposition 6. *Every accumulation point, if any, of $\{x^k\}$ is a stationary point of Problem (1).*

Proof. See [17, Theorem 3.5]. □

In what follows we present the main novelty of this paper. For the convergence of the method we need the following hypothesis.

Assumption 1. $T \neq \emptyset$, where

$$T := \{x \in C : F(x) \preceq_K F(x^k), k = 0, 1, \dots\}.$$

Even if there exists a weakly efficient solution of Problem (1), T may be an empty set. However Assumption 1 is closely related to the completeness of the $\text{Im}(F)$, and, as reported in Luc [29], the completeness of image of F ensures the existence of efficient points. Indeed, condition $T \neq \emptyset$, is assumed in order to prove the convergence of several methods for solving classical vector optimization problems; see [1, 3, 4, 15, 18, 19, 21]. Moreover, if the sequence $\{x^k\}$ has an accumulation point, then T is nonempty; see [3, 4].

Assumption 2. Each v^k of the sequence $\{v^k\}$ is scalarization compatible, i.e., exists a sequence $\{\omega^k\} \subset \text{conv}(G)$ such that

$$v^k = P_{C-x^k}(-\widehat{\beta} J_F(x^k)^T \omega^k), \quad k = 0, 1, \dots$$

As was observed in Section 2, Assumption 2 holds if $v^k = v(x^k)$, i.e., if v^k is the exact gradient projected direction at x^k . We observe that Assumption 2 was also used in Fukuda and Graña-Drummond [17] for proving the full convergence of the sequence generated by the method in the case that F is K -convex. From now on, we will assume that Assumptions 1-2 hold.

Lemma 5. *For each $\hat{x} \in T$ and $k \in \mathbb{N}$, it holds that*

$$\langle v^k, \hat{x} - x^k \rangle \geq \widehat{\beta} \langle J_F(x^k)^T \omega^k, v^k \rangle + \|v^k\|^2.$$

Proof. Take $k \in \mathbb{N}$ and $\hat{x} \in T$. As v^k is s -compatible at x^k , then there exists $\omega^k \in \text{conv}(G)$ such that, $v^k = P_{C-x^k}(-\widehat{\beta} J_F(x^k)^T \omega^k)$. Using that v^k is a projection, we have

$$\langle -\widehat{\beta} J_F(x^k)^T \omega^k - v^k, v - v^k \rangle \leq 0, \quad v \in C - x^k.$$

In particular, for $v = \hat{x} - x^k$, we obtain

$$\langle -\widehat{\beta} J_F(x^k)^T \omega^k - v^k, \hat{x} - x^k - v^k \rangle \leq 0.$$

So, from the last inequality, we get

$$\langle v^k, \hat{x} - x^k \rangle \geq -\widehat{\beta} \langle J_F(x^k)^T \omega^k, \hat{x} - x^k \rangle + \widehat{\beta} \langle J_F(x^k)^T \omega^k, v^k \rangle + \|v^k\|^2. \quad (5)$$

Since F is K -quasiconvex, and using Theorem 1, we have that $\langle d, F \rangle : \mathbb{R}^n \rightarrow \mathbb{R}$ is a quasiconvex function for each $d \in \text{extd}(K^*)$. Thus, as $\text{co}(\text{conv}(\text{extd}(K^*))) = K^*$ and $\omega^k \in \text{conv}(G) \subset K^*$, then $\omega^k = \sum_{l=1}^p \gamma_l^k d_l$, where $\gamma_l^k \in \mathbb{R}_+$ and $d_l \in \text{extd}(K^*)$, for all $1 \leq l \leq p$. Therefore,

$$\widehat{\beta} \langle J_F(x^k)^T \omega^k, \hat{x} - x^k \rangle = \widehat{\beta} \langle J_F(x^k)^T \sum_{l=1}^p \gamma_l^k d_l, \hat{x} - x^k \rangle,$$

and hence,

$$\widehat{\beta} \langle J_F(x^k)^T \omega^k, \hat{x} - x^k \rangle = \widehat{\beta} \sum_{l=1}^p \gamma_l^k \langle J_F(x^k)^T d_l, \hat{x} - x^k \rangle.$$

As $F(\hat{x}) \preceq_K F(x^k)$, we have $F(x^k) - F(\hat{x}) \in K$. So, $\langle d_l, F(x^k) - F(\hat{x}) \rangle \geq 0$, for all $d_l \in \text{extd}(K^*)$, concluding that $\langle d_l, F(\hat{x}) \rangle \leq \langle d_l, F(x^k) \rangle$. But as $\langle d_l, F \rangle$ is a real-valued, quasi-convex differentiable function, we get

$$\langle J_F(x^k)^T d_l, \hat{x} - x^k \rangle \leq 0.$$

This implies that

$$\widehat{\beta} \langle J_F(x^k)^T \omega^k, \hat{x} - x^k \rangle \leq 0.$$

Now, the result follows from combining of the last inequality with (5). \square

In next lemma we present the main result of this section. It is fundamental to the proof of the global convergence result of the sequence $\{x^k\}$.

Lemma 6. *Suppose that F is K -quasiconvex. Then, the sequence $\{x^k\}$ is quasi-Fejér convergent to the set T .*

Proof. Since $T \neq \emptyset$, take $\hat{x} \in T$ and fix $k \in \mathbb{N}$. We have the following equality:

$$\|x^{k+1} - \hat{x}\|^2 = \|x^k - \hat{x}\|^2 + \|x^{k+1} - x^k\|^2 + 2\langle x^k - x^{k+1}, \hat{x} - x^k \rangle,$$

which, from the definition of x^{k+1} , leads to

$$\|x^{k+1} - \hat{x}\|^2 = \|x^k - \hat{x}\|^2 + \|x^{k+1} - x^k\|^2 - 2t_k \langle v^k, \hat{x} - x^k \rangle. \quad (6)$$

Using Lemma 5, recall that $t_k \in (0, 1)$, we get

$$\|x^{k+1} - x^k\|^2 - 2t_k \langle v^k, \hat{x} - x^k \rangle \leq t_k \|v^k\|^2 - 2t_k \left(\widehat{\beta} \langle J_F(x^k)^T \omega^k, v^k \rangle + \|v^k\|^2 \right). \quad (7)$$

On the other hand,

$$t_k \|v^k\|^2 - 2t_k \left(\widehat{\beta} \langle J_F(x^k)^T \omega^k, v^k \rangle + \|v^k\|^2 \right) \leq -2t_k \widehat{\beta} \langle J_F(x^k)^T \omega^k, v^k \rangle,$$

and recalling that $\alpha \leq |\alpha|$, for all $\alpha \in \mathbb{R}$, from (7), we obtain

$$\|x^{k+1} - x^k\|^2 - 2t_k \langle v^k, \hat{x} - x^k \rangle \leq 2t_k |\widehat{\beta} \langle J_F(x^k)^T \omega^k, v^k \rangle|.$$

Combining last inequality with (6), we get

$$\|x^{k+1} - \hat{x}\|^2 \leq \|x^k - \hat{x}\|^2 + 2t_k \hat{\beta} |\langle J_F(x^k)^T \omega^k, v^k \rangle|. \quad (8)$$

Due to K is a pointed, closed and convex cone, we have that $\text{int}(K^*) \neq \emptyset$; see, for instance, [32, Propositions 2.1.4, 2.1.7(i)]. Therefore, the dual cone, K^* , contains a basis of \mathbb{R}^m , say $\{\tilde{\omega}^1, \dots, \tilde{\omega}^m\}$. Without loss of generality, we assume that $\{\tilde{\omega}^1, \dots, \tilde{\omega}^m\} \subset \text{conv}(G)$. Thus, for each k , there exist $\eta_i^k \in \mathbb{R}, i = 1, \dots, m$, such that

$$\omega^k = \sum_{i=1}^m \eta_i^k \tilde{\omega}^i.$$

In view of the compactness of $\text{conv}(G)$, all scalars η_i^k are uniformly bounded, which means that there exists $L > 0$, such that $|\eta_i^k| \leq L$ for all i and k . Thus, inequality (8) becomes

$$\|x^{k+1} - \hat{x}\|^2 \leq \|x^k - \hat{x}\|^2 + 2t_k \hat{\beta} L \sum_{i=1}^m |\langle \tilde{\omega}^i, J_F(x^k) v^k \rangle|.$$

Defining

$$\epsilon_k := 2t_k \hat{\beta} L \sum_{i=1}^m |\langle \tilde{\omega}^i, J_F(x^k) v^k \rangle|,$$

it follows that $\epsilon_k > 0$. Since $\{F(x^k)\}$ is K -bounded and using Lemma 3, we have $\sum_{k=0}^{\infty} \epsilon_k < \infty$. Therefore, since \hat{x} is an arbitrary element of T , the desired result follows from Definition 7. \square

Next theorem establishes a sufficient condition for the convergence of the sequence $\{x^k\}$. This result is the main convergence result in the quasiconvex case.

Theorem 2. *Assume that F is a K -quasiconvex function. Then, $\{x^k\}$ converges to a stationary point.*

Proof. Since F is K -quasiconvex, from Lemma 6 it follows that $\{x^k\}$ is quasi-Fejér convergent and, hence, bounded; see Lemma 4(i). Therefore $\{x^k\}$ has at least one accumulation point, say x^* . From Proposition 6, x^* is a stationary point. Moreover, since C is closed and the sequence is feasible, $x^* \in C$. Let $\{x^{k_j}\}$ be a subsequence of $\{x^k\}$ converging to x^* . Since F is continuous $\{F(x^k)\}$ has $F(x^*)$ as an accumulation point. By Proposition 5, $\{F(x^k)\}$ is a K -decreasing sequence. Hence, it is easy to see that the whole sequence $\{F(x^k)\}$ converges to $F(x^*)$ and the following relation holds

$$F(x^*) \preceq_K F(x^k), \quad k = 0, 1, \dots,$$

which implies that $x^* \in T$. Therefore, the desired result follows from Lemma 4 ii) and Proposition 6. \square

The last theorem extended the full convergence result presented in Fukuda and Graña-Drummond [17], before restricted to K -convex problems, to K -quasiconvex problems. In order to show that this is actually a larger class of functions, we present a K -quasiconvex problem whose vector objective function is neither K -convex nor component-wise quasiconvex.

Example 1. Let $F : \mathbb{R} \rightarrow \mathbb{R}^2$ be given by $F(t) = (f_1(t), f_2(t))$, where $f_1 := 4t^2$ and $f_2 := t^4 - 4t^2 + 2$. Take $C = [-2, 2]$ and K the generated cone by the extreme directions

$v_1 = (1, 0)$ and $v_2 = (1, 1)$. Since f_2 is not quasiconvex, then F is not component-wise quasiconvex. Moreover, F is not K -convex because, for $t_1 = 0$, $t_2 = 1$ and $\alpha = \frac{1}{2}$, we have

$$\alpha F(t_1) + (1 - \alpha)F(t_2) - F(\alpha t_1 + (1 - \alpha)t_2) = \left(1, -\frac{9}{16}\right) \notin K.$$

However, F is K -quasiconvex. Indeed,

$$\langle (1, 0), F(t) \rangle = 4t^2 \quad \text{and} \quad \langle (1, 1), F(t) \rangle = t^4 + 2,$$

are quasiconvex functions and the statement follows from Theorem 1.

5 Final Remarks

In this paper we extended the converge analysis of the sequence generated by the inexact projected gradient method presented in Fukuda and Graña-Drummond [17] to quasiconvex vectorial optimization problems. Future research is focussed in the study of a subgradient method for non-smooth K -quasiconvex functions.

Acknowledgment

The authors were partially supported by the project CAPES-MES-CUBA 226/2012, “MODELOS DE OTIMIZAÇÃO E APLICAÇÕES”. J. Y. Bello Cruz and G. C. Bento were partially supported by PROCAD-nf/CAPES, “UFG/UnB/IMPA” research and PRONEX/FAPERJ/CNPq - Optimization research. R. F. B. Costa was supported in part by CNPq.

References

- [1] J.Y. Bello Cruz, L.R. Lucambio Pérez and J.G. Melo, Convergence of the projected gradient method for quasiconvex multiobjective optimization, *Nonlinear Anal.* 74, 5268-5273 (2011).
- [2] J.Benoist, J.M. Borwein and N.Popovic, A characterization of quasiconvex vector-valued functions, *Proc. Amer. Math. Soc.* 131, 1109-1113 (2003).
- [3] G. C. Bento, O. P. Ferreira and P. R. Oliveira, Unconstrained steepest descent method for multicriteria optimization on Riemannian manifolds, *J. Optim. Theory Appl.* 154, 88-107 (2012).
- [4] G. C. Bento, J. X. Cruz Neto, P. R. Oliveira and A. Soubeyran, The self regulation problem as an inexact steepest descent method for multicriteria optimization, Ithaca: Cornell University Library, arXiv:1011.0010v1[math.OC], 2012 (Preprints).
- [5] H. Bonnel, A.N. Iusem and B.F. Svaiter, Proximal methods in vector optimization, *SIAM J. Optim.* 15, 953-970 (2005).
- [6] F.E. Browder, Convergence theorems for sequences of nonlinear operators in Banach spaces, *Math.Zeitschr* 100, 201-225 (1967).
- [7] R. Burachik, L.M. Graña Drummond, A.N. Iusem and B.F. Svaiter, Full convergence of the steepest descent method with inexact line searches, *Optimization* 32, 137-146 (1995).
- [8] E. Carrizosa and J.B.G. Frenk, Dominating sets for convex functions with some applications, *J. Optim. Theory Appl.* 96, 281-295 (1998).

- [9] L.C. Ceng, B.S. Mordukhovich and J.C. Yao, Hybrid Approximate Proximal Method with Auxiliary Variational Inequality for Vector Optimization, *J. Optim. Theory Appl.* 146, 267-303 (2010).
- [10] G.Y. Chen and J. Jahn, Optimality conditions for set-valued optimization problems. Set-valued optimization, *Math. Methods Oper. Res.* 48, 187-200 (1998).
- [11] Z. Chen, C. Xiang, K. Zhao and X. Liu, Convergence analysis of Tikhonov-type regularization algorithms for multiobjective optimization problems, *Appl. Math. Comput.* 211, 167-172 (2009).
- [12] H. Eschenauer, J. Koski and A. Osyczka, *Multicriteria Design Optimization*, Springer, Berlin (1990).
- [13] J. Fliege, OLAF - A general modeling system to evaluate and optimize the location of an air polluting facility, *OR Spectrum.* 23, 117-136 (2001).
- [14] J. Fliege, L.M. Graña Drummond and B.F. Svaiter, Newton's method for multiobjective optimization, *SIAM J. Optim.* 20, 602-662 (2009).
- [15] J. Fliege and B.F. Svaiter, Steepest descent methods for multicriteria optimization, *Math. Methods Oper. Res.* 51, 479-494 (2000).
- [16] Y. Fu and U.M. Diwekar, An efficient sampling approach to multiobjective optimization, *Annals Oper. Res.* 132, 109-134 (2004).
- [17] E.H. Fukuda and L.M. Graña Drummond, Inexact projected gradient method for vector optimization, *Comput. Optim. Appl.*, 1-21 (2012).
- [18] E.H. Fukuda and L.M. Graña Drummond, On the convergence of the projected gradient method for vector optimization, *Optimization* 60, 1009-1021 (2011).
- [19] L.M. Graña Drummond and A.N. Iusem, A projected gradient method for vector optimization problems, *Comput. Optim. Appl.* 28, 5-30 (2004).
- [20] L.M. Graña Drummond, N. Maculan and B.F. Svaiter, On the choice of parameters for the weighting method in vector optimization, *Math. Program.* 111, 201-216 (2008).
- [21] L.M. Graña Drummond and B.F. Svaiter, A steepest descent method for vector optimization, *J. Comput. Appl. Math.* 175, 395-414 (2005).
- [22] M. Gravel, J.M. Martel, R. Nadeau, W. Price and R. Tremblay, A multicriterion view of optimal resource allocation in job-shop production, *European J. Oper. Res.* 61, 230-244 (1992).
- [23] M. G. Graziano, Fuzzy cooperative behavior in response to market imperfections, *Int. J. Intell. Syst.*, 27, 108-131 (2012).
- [24] M. G. Graziano and M. Romaniello, Linear cost share equilibria and the veto power of the grand coalition, *Soc Choice Welf* 38, 269-303 (2012).
- [25] A.N. Iusem, B.F. Svaiter and M. Teboulle, Entropy-like proximal methods in convex programming, *Math. Oper. Res.* 19, 790-814 (1994).
- [26] J. Jahn, *Vector optimization - Theory, applications and extensions*, Springer, Erlangen, (2003).

- [27] J. Jahn, Mathematical vector optimization in Partially Ordered Linear Spaces, Verlag Peter D. Lang, Frankfurt, (1986).
- [28] J. Jahn, Scalarization in vector optimization, Math. Program. 29, 203- 218 (1984).
- [29] D.T. Luc, Theory of Vector Optimization. Lecture Notes in Economics and Mathematical Systems, Springer, Berlin, 319 (1989).
- [30] O.L. Mangasarian, Nonlinear Programming, McGraw-Hill, New York, (1969).
- [31] D. Prabuddha, J.B. Ghosh and C.E. Wells, On the minimization of completion time variance with a bicriteria extension, Oper. Res. 40, 1148-1155 (1992).
- [32] Y.Sawaragi, H.Nakayama and T. Tanino, Theory of Multiobjective Optimization, Academic Press Orlando, (1985).
- [33] D.J. White, Epsilon-dominating solutions in mean-variance portfolio analysis, European J. Oper. Res. 105, 457-466 (1998).